Using R in a Regulatory Environment: some FDA perspectives

Paul Schuette
Scientific Computing Coordinator
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics
Paul.Schuette@fda.hhs.gov
Disclaimer

This presentation reflects the views of the author and should not be construed to represent the FDA's views or policies.
Statistical Software Clarifying Statement

“FDA does not require use of any specific software for statistical analyses, and statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21CFR part 11]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification.

As noted in the FDA guidance, E9 Statistical Principles for Clinical Trials, ‘The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available.’ Sponsors are encouraged to consult with FDA review teams and especially with FDA statisticians regarding the choice and suitability of statistical software packages at an early stage in the product development process.”

R for Regulatory Review

How is R used for regulatory review work?

• Reviewers may opt to perform their analyses using R rather than commercial packages.
• R is used for graphics and data visualization.
• Simulations in general.
• Bayesian Methods
  – JAGS
  – Stan
• Complex, Innovative Clinical Designs (PDUFA VI)
Some R packages for Biostatistics

- survival, Therneau
- Hmisc, Harrell *et al*
- DoseFinding, Bornkamp, Pinheiro, and Bretz
- gsDesign, Anderson
- Beanz, Wang *et al*
- ORCI, Sun

IDE RStudio is used extensively at FDA.
Figure 1: Mean Change (SD) from Baseline and Treatment Difference (Xiidra – Vehicle) in Eye Dryness Score in 12-Week Studies in Patients with Dry Eye Disease

[1] Based on ANCOVA model adjusted for baseline value in Study 1, and ANCOVA model adjusted for baseline value and randomization stratification factors in Studies 2-4. All randomized and treated patients were included in the analysis and missing data were imputed using last-available data. In Study 1, one Xiidra treated subject who did not have a baseline value was excluded from analysis.
Another Product Label

R Graphic. Drug for the reduction of elevated intraocular pressure (IOP) in patients with open-angle glaucoma or ocular hypertension

https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/208254lbl.pdf

Study 304: Subjects with Baseline IOP < 25 mmHg

Visit | Rhopressa (N=186) | Timolol (N=187) | Difference (95% CI) | Rhopressa - Timolol
--- | --- | --- | --- | ---
Baseline | 22.4 | 22.4 | | |
8am | 21.1 | 21.3 | | |
10am | 20.7 | 20.7 | | |
Change From Baseline
Day 15 | -4.7 | -4.9 | 0.2 (-0.4, 0.8) | |
Day 43 | -4.6 | -4.8 | 0.3 (-0.3, 0.8) | |
Day 90 | -4.5 | -5.2 | 0.6 (0.0, 1.2) | |
Visit | Rhopressa (N=120) | Timolol (N=130) | Difference (95% CI) | Rhopressa - Timolol
--- | --- | --- | --- | ---
Baseline | 26.3 | 26.0 | | |
8am | 25.2 | 24.9 | | |
10am | 24.5 | 24.0 | | |
Change From Baseline
Day 15 | -4.7 | -5.9 | 1.2 (0.3, 2.0) | |
Day 43 | -4.3 | -4.9 | 0.6 (-0.2, 1.3) | |
Day 90 | -4.5 | -6.1 | 1.6 (0.6, 2.5) | |

This table was produced based on the observed data from all randomized subjects who did not have major protocol violations. The treatment differences and two-sided CIs for comparing Rhopressa QD vs Timolol BID 0.5% were based on Analysis of Covariance (ANCOVA) adjusted for baseline IOP.
Data Anomaly Detection

Use open source software to detect potential data problems

2. Cooperative Research and Development Agreement (CRADA) with CluePoints for detecting anomalous clinical trial sites.

Example of CRADA software output
R Shiny Apps

Internal to FDA
• Waterfall Plot
• Hepatotoxicity
• Demographics
• PRO
• DABERS

External to FDA (openFDA)
• LRT app for Adverse Event analyses
Waterfall Plot

Raw change in score from baseline for each subject at C3D1, by study arm

Waterfall plot, n=584
The Hepatotoxicity tool bolsters analysis of Drug Induced Liver Injury (DILI) through a composite visualization that includes both pre-treatment and on-treatment prevalence of ALT and BILI in terms of Hy’s Law candidate laboratory Upper Limit Normal (ULN) thresholds as well as the magnitude of these elevations normalized by respective baseline test results. This analysis is particularly useful for studies in which subjects have elevated liver enzyme test results at baseline (e.g., subjects with Chronic Hepatitis C).
The Demographic Tool provides targeted descriptive statistics and safety endpoint analysis for demographic subgroups, including age, sex, race, and ethnicity. The tool has a simple user interface that dynamically walks end-users through the process of executing the analysis. The example deals with a safety endpoint analysis.

**Demographic Tool**

<table>
<thead>
<tr>
<th>Demographic Subgroups</th>
<th>TREATMENT n/N (%)</th>
<th>PLACEBO n/N (%)</th>
<th>Risk Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERALL</td>
<td>19 / 741 (2.6)</td>
<td>33 / 751 (4.4)</td>
<td>-1.83 (-3.69, 0.03)</td>
</tr>
<tr>
<td>SEX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>17 / 459 (3.7)</td>
<td>26 / 483 (5.4)</td>
<td>-1.68 (-4.33, 0.97)</td>
</tr>
<tr>
<td>Female</td>
<td>2 / 282 (0.7)</td>
<td>7 / 268 (2.6)</td>
<td>-1.90 (-4.05, 0.24)</td>
</tr>
<tr>
<td>AGE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 65</td>
<td>16 / 631 (2.5)</td>
<td>26 / 650 (4.0)</td>
<td>-1.46 (-3.41, 0.48)</td>
</tr>
<tr>
<td>&gt;= 65</td>
<td>3 / 110 (2.7)</td>
<td>7 / 101 (6.9)</td>
<td>-4.20 (-10.02, 1.61)</td>
</tr>
<tr>
<td>RACE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>15 / 636 (2.4)</td>
<td>27 / 656 (4.1)</td>
<td>-1.76 (-3.68, 0.17)</td>
</tr>
<tr>
<td>Black</td>
<td>2 / 38 (5.3)</td>
<td>0 / 36 (0.0)</td>
<td>5.26 (1.84, 12.36)</td>
</tr>
<tr>
<td>Asian</td>
<td>2 / 12 (16.7)</td>
<td>5 / 16 (31.3)</td>
<td>-14.58 (-45.57, 16.41)</td>
</tr>
<tr>
<td>American Indian</td>
<td>O / 16 (0.0)</td>
<td>0 / 9 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Native Hawaiian</td>
<td>O / 3 (0.0)</td>
<td>0 / 4 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0 / 36 (0.0)</td>
<td>1 / 30 (3.3)</td>
<td>-3.33 (-9.76, 3.09)</td>
</tr>
<tr>
<td>Missing Race</td>
<td>O / 0 (0.0)</td>
<td>0 / 0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>ETHNICITY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>3 / 231 (1.3)</td>
<td>9 / 201 (4.5)</td>
<td>-3.18 (-6.39, 0.03)</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>16 / 510 (3.1)</td>
<td>24 / 550 (4.4)</td>
<td>-1.23 (-3.51, 1.05)</td>
</tr>
<tr>
<td>Missing Ethnic</td>
<td>O / 0 (0.0)</td>
<td>0 / 0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>REGION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>9 / 461 (2.0)</td>
<td>19 / 466 (4.1)</td>
<td>-2.12 (-4.32, 0.07)</td>
</tr>
<tr>
<td>Rest of the World</td>
<td>10 / 280 (3.6)</td>
<td>14 / 285 (4.9)</td>
<td>-1.34 (-4.66, 1.98)</td>
</tr>
<tr>
<td>Canada</td>
<td>O / 0 (0.0)</td>
<td>0 / 0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td>1 / 46 (2.2)</td>
<td>2 / 42 (4.7)</td>
<td>-2.48 (-10.05, 4.10)</td>
</tr>
<tr>
<td>Europe</td>
<td>6 / 197 (3.0)</td>
<td>3 / 202 (1.5)</td>
<td>1.56 (-1.36, 4.48)</td>
</tr>
<tr>
<td>Asia</td>
<td>3 / 37 (8.1)</td>
<td>9 / 40 (22.5)</td>
<td>-14.39 (-30.04, 1.25)</td>
</tr>
<tr>
<td>Africa</td>
<td>O / 0 (0.0)</td>
<td>0 / 0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>O / 0 (0.0)</td>
<td>0 / 0 (0.0)</td>
<td></td>
</tr>
</tbody>
</table>

Source: n/a and n/a
The X-axis is on the linear scale
Treatment is TREATMENT and Control is PLACEBO

-40 -20 0 20 - Treatment Better Control Better
**FAERS data, OpenFDA**

https://openfda.shinyapps.io/LRTest/

---

**LRT Signal Analysis for a Drug**

<table>
<thead>
<tr>
<th>M</th>
<th>Preferred Term</th>
<th>Significant?</th>
<th>LLR</th>
<th>RR</th>
<th>nij</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GASTROINTESTINAL HEMORRHAGE</td>
<td><strong>p &lt; 0.05</strong></td>
<td>5910.02</td>
<td>3.09</td>
<td>9643.00</td>
</tr>
<tr>
<td>2</td>
<td>FLUSHING</td>
<td><strong>p &lt; 0.05</strong></td>
<td>5281.32</td>
<td>3.35</td>
<td>9871.00</td>
</tr>
<tr>
<td>3</td>
<td>MYOCARDIAL INFARCTION</td>
<td><strong>p &lt; 0.05</strong></td>
<td>2906.82</td>
<td>2.26</td>
<td>11664.00</td>
</tr>
<tr>
<td>4</td>
<td>CEREBROVASCULAR ACCIDENT</td>
<td><strong>p &lt; 0.05</strong></td>
<td>1299.10</td>
<td>1.88</td>
<td>7945.00</td>
</tr>
<tr>
<td>5</td>
<td>CHEST PAIN</td>
<td><strong>p &lt; 0.05</strong></td>
<td>1098.08</td>
<td>1.78</td>
<td>7995.00</td>
</tr>
<tr>
<td>6</td>
<td>DYSPNOEA</td>
<td><strong>p &lt; 0.05</strong></td>
<td>349.52</td>
<td>1.26</td>
<td>14592.00</td>
</tr>
<tr>
<td>7</td>
<td>FALL</td>
<td><strong>p &lt; 0.05</strong></td>
<td>324.09</td>
<td>1.32</td>
<td>9201.00</td>
</tr>
<tr>
<td>8</td>
<td>ASTHENA</td>
<td><strong>p &lt; 0.05</strong></td>
<td>228.75</td>
<td>1.25</td>
<td>10141.00</td>
</tr>
<tr>
<td>9</td>
<td>PRURITUS</td>
<td><strong>p &lt; 0.05</strong></td>
<td>189.57</td>
<td>1.26</td>
<td>7683.00</td>
</tr>
<tr>
<td>10</td>
<td>DIZZINESS</td>
<td><strong>p &lt; 0.05</strong></td>
<td>158.00</td>
<td>1.18</td>
<td>12903.00</td>
</tr>
<tr>
<td>11</td>
<td>PNEUMONIA</td>
<td><strong>p &lt; 0.05</strong></td>
<td>105.47</td>
<td>1.23</td>
<td>7564.00</td>
</tr>
<tr>
<td>12</td>
<td>PAIN IN EXTREMITY</td>
<td><strong>p &lt; 0.05</strong></td>
<td>5.72</td>
<td>1.04</td>
<td>7155.00</td>
</tr>
<tr>
<td>13</td>
<td>DIARRHEA</td>
<td><strong>NS</strong></td>
<td>1.10</td>
<td>1.01</td>
<td>12139.00</td>
</tr>
<tr>
<td>14</td>
<td>DRUG IN EffectIVE</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.59</td>
<td>13554.00</td>
</tr>
<tr>
<td>15</td>
<td>FATIGUE</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.95</td>
<td>14166.00</td>
</tr>
<tr>
<td>16</td>
<td>HEADACHE</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.80</td>
<td>18951.00</td>
</tr>
<tr>
<td>17</td>
<td>NAUSEA</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.87</td>
<td>14803.00</td>
</tr>
<tr>
<td>18</td>
<td>Other</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.87</td>
<td>80775.00</td>
</tr>
<tr>
<td>19</td>
<td>PAIN</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.78</td>
<td>8000.00</td>
</tr>
<tr>
<td>20</td>
<td>VOMITING</td>
<td><strong>NS</strong></td>
<td>0.00</td>
<td>0.89</td>
<td>8817.00</td>
</tr>
</tbody>
</table>

---

**Likelihood Ratio Test (LRT) Methodology**

The RR is known as the ratio of reporting rate for a particular AE for a specified drug/drug class relative to the reporting rate for all other AEs for the specified drug/drug group. RR > 1 suggests that the observed reporting rate for the particular AE is higher than the reporting rate for other AEs for the drug/drug group. An AE with RR > 1 can be considered unusual, or unusual relative to other AEs in the specified drug/drug group. RR = (a/b+c)/(a+c) (See Table 2 in Likelihood Ratio Test (LRT) Methodology document for other definitions.) Log(R) (LLR) represents the logarithm of the likelihood ratio test statistic by AEs expressed in terms of SKT, PT, etc. The larger the Log(R) value, the stronger is the association between the particular AE and (fixed) drug. Log(R) = a x log(a) - log(b) - log(c) + log(c+d) - (a+c) x log(a+c) - log(b+c) - log(d) is calculated using LogLR. AE represents the significance of the observed association between the AE and a fixed drug group. P-values less than 0.05 are indicative of those AEs being signals for the (fixed) drug. Users can use different thresholds for the p-values for signal detection (such as 0.025, 0.01, etc.).
Text Plot from LRT app,
Drug: aspirin

Text Plot for Terms. Draw a box around terms to see more details
Birthdate Problem

Basic Birth Date Problem. What is the probability that at least two subjects in a group share the same date of birth (month, day and year)?

For probability level 0.5, the required number of subjects is N = 186
R for Research

- Data Mining and Machine Learning (also with Python)
- Simulations
- Evaluation of methodology
- Oak Ridge Institute for Science and Education (ORISE) Internships
- Broad Agency Agreements (BAA)
- Cooperative Research and Development Agreements (CRADA)
- PhUSE, DIA, and ASA working groups
Research, Pediatric vs Adult ADRs

Adverse Event

Legend:
- RD = Risk in pediatric patients - Risk in adult patients

- RD < -10%
- RD -10% to -2%
- RD -5% to 0%
- RD 0% to 5%
- RD 5% to 10%
- RD > 10%

Examples of adverse events:
- Abdominal pain
- Constipation
- Dry mouth
- Dyspepsia
- Nausea
- Oral hypoglycemia
- Vomiting
- Additional injury
- Edema
- Fatigue
- Inflicted injury
- Pain
- Respiratory tract infection
- Hepatic enzyme increased
- Triglycerides increased
- Weight increased
- Weight decreased
- Increased appetite
- Dizziness
- Extrapyramidal disorder
- Headache
- Sedation
- Agitation
- Anxiety
- Insomnia
- Menstrual disorders
- Discontinuation
- Serious adverse event
Concluding Observations

• Open source tools such as R offer cost effective ways for FDA to carry out its public health mission, and to enhance communications with the public, health care providers and regulated industry.

• R is widely used in academe, and is the first choice for many recent graduates.

• Managing packages and dependencies can be challenging.

• Interactive tools such as R Shiny can enhance users’ experience and understanding.

• We still need subject matter experts to help frame questions and draw appropriate conclusions.
Acknowledgements

• Yan Wang
• Jimmy Wong
• Daniel Choi
• Austin Taylor
• Xiaomei Liu